DOCUMENT RESUME

ED 057 076 TM 000 913

AUTHOR	Federico, Pat-Anthony
TITLE	Identifying Item Validity Indices Utilizing a Multivariate Model.
INSTITUTION	Air Force Human Resources Lab., Brooks AFB, Texas.
REPORT NO	AFHRL-TR-71-16
PUB DATE	Apr 71
NOTE	16p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Attitude Tests; *Discriminant Analysis; *Item Analysis; Predictive Ability (Testing); Psychometrics; Statistical Analysis; Student Attitudes; *Surveys; Technical Education; *Validity; *Weighted Scores

ABSTRACT

This study demonstrates and discusses a new procedure for performing item analysis which utilizes multiple discriminant analysis to establish efficiently and effectively an index of item validity. Application of this statistical technique to data derived from an attitude survey of three groups of students enrolled in technical training courses yielded the following results: It disclosed those stimulus items which were responsive enough to discriminate among criterion groups; it partitioned the total discriminatory power of the items into two homogeneous components; it yielded data for arriving at a special weighting scheme for scoring the final attitude form; and it located the positions of the criterion groups relative to the two orthogonal dimensions of the attitude universe. (Author)

AIR FORCE

HUMAN RESOURCES LABORATORY

ED057076

AFHRL-TR-71-16

IDENTIFYING ITEM VALIDITY INDICES UTILIZING A MULTIVARIATE MODEL

By

Pat-Anthony Federico, 1st Lt, USAF

TECHNICAL TRAINING DIVISION
Lowry Air Force Base, Colorado

April 1971

Approved for public release; distribution unlimited.

AIR FORCE SYSTEMS COMMAND

BROOKS AIR FORCE BASE, TEXAS

NOTICE

When US Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

AFHRL-TR-71-16

April 1971

IDENTIFYING ITEM VALIDITY INDICES UTILIZING A MULTIVARIATE MODEL

By

Pat-Anthony Federico, 1st Lt, USAF

TECHNICAL TRAINING DIVISION
AIR FORCE HUMAN RESOURCES LABORATORY
AIR FORCE SYSTEMS COMMAND
Lowry Air Force Base, Colorado

FOREWORD

This study represents a portion of the in-house research program of Project 1121, Technical Training Development; Task 112103, Evaluating Individual Proficiency and Technical Training Programs. Dr. Marty R. Rockway was the project scientist and Capt Wayne S. Sellman was the task scientist. This report covers research performed during February 1971 and April 1971.

The author wishes to express appreciation to 1st Lt Edward M. Gardner and Mr. Gerald S. Walker for their assistance in debugging and running statistical routines from the Laboratory's computer library.

This report has been reviewed and is approved.

George K. Patterson, Colonel, USAF
Commander

4

ABSTRACT

This study demonstrates and discusses a new procedure for performing item analysis which utilizes multiple discriminant analysis to establish efficiently and effectively an index of item validity. Application of this statistical technique to data derived from an attitude survey of three groups of students enrolled in technical training courses yielded the following results: It disclosed those stimulus items which were responsive enough to discriminate among criterion groups; it partitioned the total discriminatory power of the items into two homogeneous components; it yielded data for arriving at a special weighting scheme for scoring the final attitude form; and it located the positions of the criterion groups relative to the two orthogonal dimensions of the attitude universe.

SUMMARY

Federico, Pat-Anthony. *Identifying item validity indices utilizing a multivariate model.* AFHRL-TR-71-16. Lowry AFB, Colo.: Technical Training Division. Air Force Human Resources Laboratory, April 1971.

Problem

An index of item validity is typically computed to ascertain how well an item measures or discriminates in agreement with the rest of a test, or how well an item predicts some external criterion. Like other item analysis techniques, it is used in the selection of the best items from which to compose a final test or attitude form. The purpose of this study was to introduce a new method of determining an index of item validity: specifically, an index which can be established efficiently and effectively by the utilization of multiple discriminant analysis (DSCRIM).

Approach

As part of a task to identify valid and reliable psychometric measures of student attitudes towards Air Force technical training, DSCRIM was performed on data derived from an attitude survey of three groups of trainees enrolled in courses at the Technical Training Center, Lowry Air Force Base, Colorado.

Results and Conclusions

It was demonstrated that DSCRIM could extract many different kinds of information from data which normally would have been obtained from the execution of numerous item analysis techniques. Namely, DSCRIM accomplished the following results: It disclosed those stimulus items which were responsive enough to discriminate among criterion groups; it partitioned the total discriminatory power of the items into two homogeneous components; it yielded data for arriving at a special weighting scheme for scoring the final form; and it located the positions of the criterion groups relative to the two orthogonal dimensions of the attitude universe. Not only did DSCRIM establish several distinct item validity indices, but also it did so more efficiently (it derived these different indices simultaneously with minimum effort) and more effectively (it practically exhausted the information obtainable from the data concerning item selection) than other item analysis procedures.

This summary was prepared by Pat-Anthony Federico, Technical Training Division, Air Force Human Resources Laboratory.

TABLE OF CONTENTS

LIST OF TABLES

FIGURE

7

IDENTIFYING ITEM VALIDITY INDICES UTILIZING A MULTIVARIATE MODEL

I. INTRODUCTION

According to Guilford (1954), item analysis is typically used in the selection of the best items from which to compose a final test form. Starting with a surplus number of items, the writer can retain those which meet certain criteria specified by several statistics. Item analysis is performed to establish indices of item difficulty and item validity. As an index of item difficulty, it determines how troublesome an item was for a particular population. As an index of item validity, it determines how well an item measures or discriminates in agreement with the rest of the test, or how well an item predicts some external criterion.

The most common statistics computed for the item difficulty index are the empirical probability that the population involved will pass a particular item (Coombs, 1950); or the correlation of an item with an internal criterion (usually the total score) obtained from only an extreme part of the population (Michael, Hertzka, & Perry, 1953). Other statistics that have been proposed as indices of item difficulty have been described by Gulliksen (1950). The most common procedures followed for establishing the item validity index are the division of the total sample into groups according to a criterion and the determination of whether these groups behave differently with respect to an item (Lawshe, 1942); the subdivision of the sample on the basis of total scores into seven groups at intervals of $.6s_t$ and the calculation of the proportion of each subgroup who pass the item (Ferguson, 1942); the correlation of an item with an external criterion score in some way (Guilford, 1954); and the implementation of analysis of variance techniques which are claimed to extract the utmost information for item analysis (Baker, 1939).

The purpose of this article is to demonstrate a new method of conducting item analysis: an index of item validity that can be established more efficiently and effectively than those methods mentioned by the utilization of multiple discriminant analysis. This is a statistical routine for determining a linear combination of p variables which, better than any other linear combination, discriminates among n groups. Specifically, the ratio of the among-groups sum-of-squares of this linear combination to its within-groups sum-of-squares (the discriminant criterion) has a larger value than that for any other linear combination of the same variables. This optimum linear combination is called the discriminant function. Considering this combination as the one and only discriminant function, in effect, makes a linear ordering of the n groups. Consequently, the discriminant function thus defined does not necessarily utilize all the information in the data relevant to group separation, except when the group centroids (multivariate means) are in fact colinear. It is likely that when n is greater than 2, more than one discriminant function may be necessary to exhaust the information in the data relevant to group separation. The first linear combination maximizes the discriminant criterion; the second linear combination maximizes the ratio of the residual among-groups sum-of-squares after the effect of the first linear combination has been removed; the third linear combination maximizes the ratio of the corresponding sums-of-squares after the effect of the first two have been removed; and so forth. These successive linear combinations are referred to as multiple discriminant functions; these functions are all considered to be mutually orthogonal. For n groups there are n-1 of these, except in the unusual case when the number of variables, p, is smaller then n-1; in that case there are p discriminant functions (Fisher, 1938; Bryan, 1951; Tiedeman, 1951; Rao, 1952; Tatsuoka & Tiedeman, 1954; Wert, Neidt, & Ahmann, 1954; Cooley & Lohnes, 1966).

II. METHOD

Subjects

Subjects were 100 officers, 90 NCOs, and 99 airmen who were taking courses at the Technical Training Center, Lowry Air Force Base, Colorado. Entire classes were randomly selected from five different technical departments; all *S*s had been enrolled in their respective departments for one month or more. This period of time was assumed to be long enough for them to have crystallized some attitudes toward Air Force technical training.

Attitude Form Construction

As part of a project to identify valid and reliable psychometric measures of student attitudes toward Air Force technical training, the contents of the stimulus items used in this investigation were generated by random samples of students. The content of each of 55 items was structured according to a Likert format (Likert, 1932; Edwards, 1957); each item was randomly placed in its sequential position among the 55 items. In typical Likert fashion, the items called for checking one of five responses: strongly agree, agree, undecided, disagree, and strongly disagree. These five response categories were scored 5, 4, 3, 2, and 1, respectively, for items favorable to Air Force technical training; the scoring of the responses was reversed for items unfavorable to Air Force technical training. This procedure of scoring by assigning arbitrary weights to response categories correlates highly with normal deviate weighting of response categories. Also, using this procedure produces no noticeable differences in reliabilities (Likert, 1932). The total score for a *S* was determined by summing his responses to all the items. In order to avoid any position bias in responding, the agreeable end of the response continuum was alternated randomly from the left to the right among the items.

Procedure

The attitude form was presented to the *S*s in booklet form on the cover of which appeared generalized information descriptive of the task. The following page contained a questionnaire which concerned itself with some aspects of the *S*'s personal history. Next in sequence was the attitude form, preceded by its specific instructions. Following this form was an open-ended questionnaire designed to give the *S*s an opportunity to comment on various aspects of the form and the experimental situation. The booklets were distributed to the *S*s in their actual classrooms; typically, classes consisted of approximately 12 students. The booklets were considered to be self-explanatory since they contained all the necessary instructions. While *S*s were responding to the attitude form and the questionnaires, they were supervised by an assistant in order to preclude inter-*S* collaboration or response contamination.

III. RESULTS

Assuming that the three samples of *S*s were drawn from three populations in which the joint distributions of the scored responses to the 55 stimulus items were multivariate normal with equal dispersions, a multiple discriminant analysis (DSCRIM) was performed on the data using Veldman's (1967) computer program. The resulting Wilks' lambda criterion (λ) for the discriminating power of the attitude items to separate the three groups indicated the chance was essentially zero that group differences as large as or larger than those obtained would be produced by drawing three random samples from a 55 dimensional multivariate swarm. In Table 1 it can be seen that 100 percent of the trace was accounted for by two discriminant functions. The eigenvectors presented in Table 2 are the coefficients of the discriminant functions. These discriminant weights, together with the correlations listed in Table 3, indicated the consequential contributors to group separation along the first and second functions.

The aforementioned statistics suggested the following interpretations. Firstly, the original attitude form was sensitive enough to separate significantly three groups of Air Force students who had been known to have drastically different attitudes towards military training in general, namely, officers, NCOs, and airmen. Secondly, the total discriminatory power of the items was accounted for by two mutually orthogonal functions. The relative sizes of their eigenvalues established the extent to which the discriminant functions distinguished among the groups. The degree of importance attached to each of these functions in explaining the differences among the groups had been estimated by the percentage of the trace attributed to each of these functions. Consequently, the first function was considered to be more important than the second function for simplifying group separation. Thirdly, the relative contributions of the items to the discriminant functions had been determined by the magnitude of the correlations between the items and the discriminant functions. These coefficients had been interpreted like factor loadings so that the two functions were described and named in the context of the attitude items having significant correlations with the functions. Therefore, the

Table 1. **Significance of the Discriminant Functions χ^2 Approximations**

Function	Percent of Trace	Eigen-values	df	χ^2	p
I	68.20	1.884	56	273.267	0.000
II	31.80	.878	54	162.664	0.000

Trace = 2.762; $\lambda = 0.185$; $F(110,460) = 5.552$; $p = 0.000$

9

Table 2. **Discriminant-Function Weights for Each Stimulus Item**

Item	Function I	Function II	Item	Function I	Function II
1	−0.0390	−0.0746	29	−0.1873	−0.1129
2	−0.0869	−0.1260	30	−0.1228	0.0191
3	−0.0057	0.1823	31	−0.0520	−0.0971
4	0.2117	0.1560	32	−0.0844	−0.1315
5	−0.1240	−0.1027	33	0.2022	0.0202
6	−0.0486	0.1134	34	0.1106	0.0895
7	−0.0385	0.0808	35	−0.0639	0.0008
8	−0.0942	0.1516	36	0.0257	−0.2038
9	0.0448	0.1204	37	0.2541	0.1313
10	0.2148	−0.1339	38	−0.2126	−0.0905
11	−0.1225	−0.1138	39	−0.1700	0.0383
12	0.1724	0.0184	40	−0.0072	0.0361
13	0.1555	−0.0194	41	0.0834	0.1305
14	0.0909	−0.1835	42	−0.0825	−0.1056
15	0.0595	0.1691	43	0.1078	0.1945
16	−0.0583	−0.0045	44	0.0516	0.1431
17	0.0707	−0.0004	45	−0.0454	−0.1320
18	0.0219	−0.0638	46	−0.0892	−0.1583
19	−0.0305	0.1110	47	0.1823	−0.1022
20	0.3049	0.0410	48	−0.0756	−0.0155
21	0.0091	−0.0364	49	0.0814	0.0375
22	0.1322	−0.0912	50	−0.2874	0.1890
23	−0.0231	0.0427	51	−0.0565	0.0729
24	−0.0601	0.1483	52	0.0395	−0.0972
25	−0.0701	0.2030	53	0.0827	0.1536
26	−0.0909	−0.3985	54	0.1846	0.0521
27	−0.0444	0.0455	55	0.1558	0.1603
28	−0.3960	0.3759			

Table 3. **Correlations Between Stimulus Items and the Discriminant Functions**

Item	Function I	Function II	Item	Function I	Function II
1	0.2450*	0.0225	29	0.0060	0.0917
2	0.1434	−0.0632	30	0.1208	−0.0277
3	−0.0940	0.4503*	31	0.1468	0.0648
4	0.4747*	0.2417*	32	−0.0114	0.0094
5	−0.1787	0.1936*	33	0.2287*	0.2183*
6	0.0349	0.2056*	34	0.0847	0.0783
7	0.2577*	0.1505*	35	0.0753	0.0116
8	0.0220	0.1109	36	0.0915	−0.0543
9	0.1277	−0.0239	37	0.4616*	0.0807
10	0.2411*	0.0317	38	−0.1965*	−0.0228
11	−0.1143	0.0120	39	−0.0288	0.0974
12	0.1503*	0.1085	40	0.2279*	0.0752
13	0.3984*	0.0301	41	−0.0837	0.3024*
14	0.1156	0.0212	42	−0.1638*	0.0163
15	0.0548	0.2415*	43	0.2114*	0.1184
16	−0.1526*	0.1055	44	0.0950	0.2584*
17	−0.0039	0.1581*	45	0.2934*	−0.1828*
18	0.2226*	−0.0198	46	0.1506*	−0.2132*
19	−0.0965	0.2110*	47	0.3801*	0.0053
20	0.6347*	0.1113	48	−0.0615	0.3093*
21	0.0803	0.0578	49	0.0595	0.3236*
22	0.2112*	−0.1434	50	−0.2046*	0.2800*
23	0.0045	0.1095	51	−0.1212	0.1240
24	−0.1026	0.1816*	52	0.3245*	0.2385*
25	0.0207	0.1445	53	0.3788*	0.5010*
26	0.0444	−0.0551	54	0.2448*	0.1745*
27	0.0045	0.2112*	55	0.4434*	0.3659*
28	−0.3708*	0.6003*			

*$r \geqslant 0.148$ is significant at the .01 level, $N = 287$.

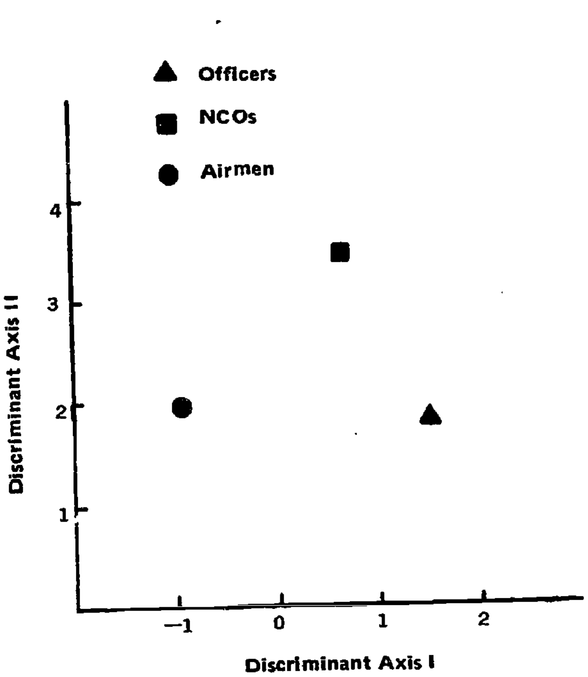

Fig. 1. **Group centroids in the discriminant space.**

first and second discriminant functions were labelled Training Management and Training Impressions, respectively. Lastly, another result of multiple discriminant analysis is that the group centroids computed relative to the functions are separated from each other to a maximum degree. Considering the two discriminant functions as axes of a Cartesian coordinate system, the group centroids were plotted with reference to these axes in Figure 1. It can be seen that the officers sampled had manifested more favorable attitudes towards training management than the NCOs, and the airmen sampled had manifested more unfavorable attitudes towards training managment than either the officers or the NCOs. Also depicted is the fact that the NCOs had demonstrated more favorable training impressions than the officers and the airmen. Surely, these final findings hardly seem surprising to soldiers, past and present.

IV. DISCUSSION

DSCRIM identified those items included in the original attitude form which were sensitive enough to distinguish among the preseparated groups. It demonstrated that the total discriminatory power of these items was accounted for by two orthogonal functions. It established the relative contributions of the items to each discriminant function. Also, it ascertained the positions of the groups with respect to the discriminant axes. These distinct determinations, derived from one DSCRIM, would have required the implementation of different item analysis procedures in conjunction; namely, item validity indices which demonstrate how well an item discriminates with the rest of the attitude form, correlates with some internal or external criterion, and weights several unidimensional continua. DSCRIM was capable of identifying these indices, plus designating the positions of preseparated groups relative to the orthogonal axes along which they differed, all in one computer run.

Partitioning of the discriminatory power of the original set of attitude items into independent components reduced item space dimensionality without substantial loss of information. The aforementioned techniques of item analysis which are currently used to estimate item validity indices do not even establish dimensionality, let alone reduction of dimensionality. DSCRIM not only determined the orthogonal axes of the item space, but also it determined how well items differentiated along each axis. Knowledge of the relationship of an item to an orthogonal function which distinguished preseparated groups yielded information concerning the number and kind of discriminations the item was capable of making. DSCRIM emphasized the extent to which items separated the criterion groups along the derived dimensions of the attitude space. In addition to denoting which of the original items were responsive enough to differentiate among the three criterion groups, DSCRIM defined orthogonal continua which underlay the discriminations among the group attitudes.

When the original attitude form was tried out, it included many items which had only an ostensible relationship to one another. Initially, items were not held together by any underlying continua, but rather by the superficial fact that all the items referred to statements about likes and dislikes regarding the attitude universe. There might have been very little uniformity in terms of what the individual items purported to measure. Convenience and comparable contents of the items might have made it desirable to include all of them in the final form. There was no reason, though, to think of all the items as constituting a homogeneous attitude form measuring the same common factor in all individuals. To say the least, it is nice to know when constructing an attitude form or aptitude test, to what degree a particular item measures the same orthogonal dimension as other items. After all, a scale score is meaningful only to the extent that each item contributes to the measurement of a single homogeneous function. By determining the discriminant axes along which the criterion groups differed, DSCRIM also established homogeneous clusters of items, like factor analysis. The correlations among items and discriminant functions were interpreted in a manner similar to factor loadings. Those items which significantly weighted these independent components were retained for inclusion within the final form. Not only did DSCRIM disclose homogeneous dimensions, but also it designated simultaneously those dimensions with respect to which items operated as discriminators to maximize distance among criterion groups.

The validity of a scale score may be substantially increased by the optimal weighting of each of its composite items. According to Thorndike (1969), the only justification for arriving at a special weighting scheme is in terms of the empirical validity of a cluster of items. The effectiveness of the items in discriminating with respect to some external criterion must be a paramount consideration before assigning item weights. DSCRIM provided information regarding the differential weighting of items for scoring the final attitude form. Not all items were found to have significant correlations with each of the discriminant functions which maximized the separation among criterion groups. The proposed DSCRIM procedure for weighting items has a distinct advantage over the multiple regression procedure for weighting items. An inherent characteristic of DSCRIM is that it maximizes intergroup distances by optimally weighting linear combinations of individual items. Therefore, scoring weights derived from DSCRIM are more susceptible to detecting attitudinal differences along the independent dimensions of criterion discrimination than are scoring weights derived from multiple regression analysis.

11

V. SUMMARY AND CONCLUSIONS

It can be seen that DSCRIM extracted many different kinds of information from the data which normally would have been obtained from the implementation of numerous item analysis techniques. To recapitulate, DSCRIM achieved the following results: It disclosed those stimulus items which were responsive enough to discriminate among criterion groups; it partitioned the total discriminatory power of the items into two homogeneous components; it yielded data for arriving at a special weighting scheme for scoring the final form; and it located the positions of the criterion groups relative to the two orthogonal dimensions of the attitude universe. Essentially, DSCRIM produced useful information regarding the selection of the best items from which to compose the final attitude form. Not only did DSCRIM establish several distinct item validity indices, but also it did so more efficiently (it derived these different indices simultaneously with minimum effort) and more effectively (it practically exhausted the information obtainable from the data concerning item selection) than other item analysis procedures.

REFERENCES

Baker, K.H. Item validity by the analysis of variance: An outline of method. *Psychological Record,* 1939, **3**, 242-248.

Bryan, J.G. The generalized discriminant function: Mathematical foundation and computational routine. *Harvard Educational Review,* 1951, **21** (2), 90-95.

Cooley, W.W. & Lohnes, P.R. *Multivariate procedures for the behavioral sciences.* New York: John Wiley, 1966.

Coombs, C.H. The concepts of reliability and homogeneity. *Journal of Educational and Psychological Measurement,* 1950, **10**, 43-56.

Edwards, A. *Techniques of attitude scale construction.* New York: Appleton-Century-Crofts, 1957.

Ferguson, G.A. Item selection by the constant process. *Psychometrika,* 1942, **7**, 19-29.

Fisher, R.A. The statistical utilization of multiple measurements. *Annals of Eugenics,* 1938, **8**, 376-386.

Guilford, J.P. *Psychometric methods.* New York: McGraw-Hill, 1954.

Gulliksen, H. *Theory of mental tests.* New York: John Wiley, 1950.

Lawshe, C.H. A nomograph for estimating the validity of test items. *Journal of Applied Psychology,* 1942, **26**, 846-849.

Likert, R.A. A technique for the measurement of attitudes. *Archives of Psychology,* 1932, **22** (140), 1-55.

Michael, W.B., Hertzka, A.F., & Perry, N.C. Errors in estimates of item difficulty obtained from the use of extreme groups on a criterion variable. *Journal of Educational and Psychological Measurement,* 1953, **13**, 601-606.

Rao, C.R. *Advanced statistical methods in biometric research.* New York: John Wiley and Sons, 1952.

Tatsuoka, M.M. & Tiedeman, D.V. Discriminant analysis. *Review of Educational Research,* 1954, **24** (5), 402-420.

Thorndike, R.L. *Personnel selection: Test and measurement techniques.* New York: John Wiley, 1969.

Tiedeman, D.V. The utility of the discriminant function in psychological and guidance investigations. *Harvard Educational Review,* 1951, **21** (2), 71-80.

Veldman, D. *Fortran programming for the behavioral sciences.* New York: Holt, Rinehart, and Winston, 1967.

Wert, J.E., Neidt, C.O., & Ahmann, J.S. *Statistical methods in educational and psychological research.* New York: Appleton-Century-Crofts, 1954.

ERRATA

Huff, K.H. & Smith, E.A. *Reliability, baseline data, and instructions for the automated readability index.* Lowry Air Force Base, Colo.: Technical Training Division, Air Force Human Resources Laboratory, October 1970. (AFHRL-TR-70-14)

Substitute the attached page for page 7.

APPENDIX: TYPING AND COMPUTATIONAL PROCEDURES FOR THE AUTOMATED READABILITY INDEX

For effective use of the ARI method of measuring readability, the equipment must be accurate and the typing procedures correct. This appendix contains a set of general instructions and a practice sentence and paragraph for which the number of strokes, words, and sentences are known. A typist using the method should practice the modified typing procedure and check the apparatus before proceeding, as well as make periodic checks on the apparatus throughout the data tabulation and collection phase. The need for such instruction was confirmed by a feasibility test of the Automated Readability Index apparatus conducted by Air Training Command (ATC Project Report 69-22).

General Instructions for Computing the Automated Readability Index

Selecting Samples

1. Only books and passages containing straight narrative material can be used. Skip unusual material such as poems, listings, math, etc.

2. For each book, the equivalent of 10 full pages comprise the sample. Start at the beginning of a paragraph and end at the ending of a paragraph - this means that the material copied will not physically appear on one page in the text, but will be approximately one page in length.

3. For each page typed, indicate the source and page. On the data sheet of each source, give all of the information needed for citing in a bibliography. Typed pages need only enough to identify source and page within source.

4. Select 10 pages, spaced relatively evenly through the book. Try to avoid the first and last page of chapters. For example, if the book has 325 pages, type page 16, then every 32nd page. However, alternate between left and right hand pages (even and odd). Don't be overly compulsive and waste undue time deciding exactly which page to use - in general get 10 pages that will represent the full book, remembering that some books change markedly as you progress through, i.e., the latter protions are much more difficult than the beginning.

Typing-Instructions

1. Very few changes from standard typing are required. Use common sense to formulate rules. You must end with a count of

    a. the number of letters on the page.
    b. the number of words on the page.
    c. the number of sentences on the page.

2. Getting the number of letters is fairly simple if you remember that such things as punctuation marks are included in the count. They should be counted. However, notice that the spelling of a word is immaterial, as long as it contains the proper number of letters. Please don't make corrections or strike-overs. REMEMBER YOUR TASK IS TO OBTAIN COUNTS AND ONLY THE COUNTING IS IMPORTANT. The tabulator will register each time you type a letter.

3. Counting the words is done by counting the number of times the space bar is used. This requires a few simple changes in typing.

    a. Do not double space after a sentence - to do so would add one to the word count.
    b. Do not use the space bar to indent.
    c. IMPORTANT - You must space once at the end of each typed line in order to count the last word. It would be helpful to space and then put a 7 after the last word in each typed line, except when a period occurs. This will insure the fact that you have counted the last word. Simply subtract the number 7's from the final word count.
    d. Count the numbers that enumerate items within a sentence, as words, e.g., 1. . . . , 2. . . . , 3. . .

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Technical Training Division Lowry Air Force Base, Colorado 80230	2a. REPORT SECURITY CLASSIFICATION 2b. GROUP
3. REPORT TITLE IDENTIFYING ITEM VALIDITY INDICES UTILIZING A MULTIVARIATE MODEL	
4. DESCRIPTIVE NOTES (Type of report and inclusive dates)	
5. AUTHOR(S) (First name, middle initial, last name) Pat-Anthony Federico	
6. REPORT DATE April 1971	7a. TOTAL NO. OF PAGES: 5 7b. NO. OF REFS: 18
8a. CONTRACT OR GRANT NO. b. PROJECT NO. 1121 c. Task No. 112103 d. Work Unit No. 112103002	9a. ORIGINATOR'S REPORT NUMBER(S) AFHRL-TR-71-16 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
10. DISTRIBUTION STATEMENT Approved for public release; distribution unlimited.	
11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Technical Training Division Lowry Air Force Base, Colorado 80230

13. ABSTRACT

This study demonstrates and discusses a new procedure for performing item analysis which utilizes multiple discriminant analysis to establish efficiently and effectively an index of item validity. Application of this statistical technique to data derived from an attitude survey of three groups of students enrolled in technical training courses yielded the following results: It disclosed those stimulus items which were responsive enough to discriminate among criterion groups; it partitioned the total discriminatory power of the items into two homogeneous components; it yielded data for arriving at a special weighting scheme for scoring the final attitude form; and it located the positions of the criterion groups relative to the two orthogonal dimensions of the attitude universe.

DD FORM 1 NOV 65 1473

Unclassified

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
item analysis item discrimination item validity indices item selection item weighting psychometric methods psychological testing test construction multivariate methods multiple discriminant analysis criterion group discrimination						